Exploring the binding mechanism and accessible angle of SARS-CoV-2 spike and ACE2 by molecular dynamics simulation and free energy calculation

Cheng Peng[1,2], Zhengdan Zhu[1,2], Yulong Shi[1,2], Xiaoyu Wang[1], Kaijie Mu[1,3], Yanqing Yang[1,2], Xinben Zhang[1], Zhijian Xu[1,2,*], Weiliang Zhu[1,2,*]

[1]CAS Key Laboratory of Receptor Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China

[2]School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, PR China

[3]Nano Science and Technology Institute, University of Science and Technology of China, Suzhou, Jiangsu, 215123, China.

[*]To whom correspondence should be addressed.

Phone: +86-21-50806600-1201 (Z.X.), +86-21-50805020 (W.Z.),

E-mail: zjxu@simm.ac.cn (Z.X.), wlzhu@simm.ac.cn (W.Z.).

ORCID:

Zhijian Xu: 0000-0002-3063-8473

Weiliang Zhu: 0000-0001-6699-5299

Abstract

The SARS-CoV-2 has caused more than 2,000 deaths as of 20 February 2020 worldwide but there is no approved effective drug. The SARS-CoV-2 spike (S) glycoprotein is a key drug target due to its indispensable function for viral infection and fusion with ACE2 as a receptor. To facilitate the drug discovery and development with S protein as drug target, various computational techniques were used in this study to evaluate the binding mechanisms between S protein and its acceptor ACE2. Impressively, SARS-CoV-2 S protein has higher affinity binding to ACE2 at two different "up" angles of RBD than SARS-CoV S protein to ACE2 at the same angles. The energy decomposition analysis showed that more interactions formed between SARS-CoV-2 S protein and ACE2, which may partially account for its higher infectiousness than SARS-CoV. In addition, we found that 52.2° is a starting accessible "up" angle of the BRD of SARS-CoV-2 S protein to bind ACE2, demonstrating that BRD is not necessary to be fully opened in order to bind ACE2. We hope that this work will be helpful for the design of effective SARS-CoV-2 S protein inhibitors to address the ongoing public health crisis.

1. Introduction

Very recently, a new coronavirus that is closely related to severe acute respiratory syndrome coronavirus (SARS-CoV),[1-3] temporally named SARS-CoV-2 by the international committee on taxonomy of viruses (ICTV), has emerged as a human pathogen in Wuhan, Hubei Province, China, and rapidly spread worldwide. It has caused more than 2,000 deaths as of 20 February 2020 worldwide, mostly in China, and the number is still growing. However, there is no drug has been approved to be effective. Therefore, it is very urgent to discover and develop safe and effective therapeutics.

Compared to SARS-CoV, SARS-CoV-2 is more likely to transmit from human-to-human.[4-5] The spike (S) glycoprotein of SARS-CoV-2 is a class I viral fusion protein, which plays a vital role in the viral infection with human angiotensin-converting enzyme 2 (ACE2) as a receptor, and mediating fusion of the SARS-CoV-2 and cellular membranes.[6-7] The S protein consists of an amino (N)-terminal S1 subunit and a carboxyl (C)-terminal S2 subunit. In order to recognize the ACE2, the receptor-binding domain (RBD) of S1 subunit undergoes hinge-like conformational changes to expose enough space for receptor binding.[8-10] Therefore, there are two states of S protein that are referred to as "down" and "up" conformation, where "down" conformation is the receptor-inaccessible state and "up" conformation is the receptor-accessible state.[11-14] The significant function of the S protein makes it a vital target for the drug discovery and development of the SARS-CoV-2.

In order to make a thorough understanding of the binding mechanisms between SARS-CoV-2 S protein and ACE2, various computational techniques, including MD

simulation, MM/GBSA, binding free energy decomposition analysis, and normal mode analysis (NMA) were carried out in the present study. The results not only revealed that SARS-CoV-2 S protein binds to ACE2 with higher affinity compared with SARS-CoV, even though the RBD domain is flexible with different "up" angles, but also predicted key residues of SARS-CoV-2 S protein for binding to ACE2. In addition, we found that 52.2° is an ACE2-accessible RBD "up" angle during the "down" to "up" conformational change of SARS-CoV-2 S protein. Knowledge of the interactions between SARS-CoV-2 S protein and ACE2 is required to understand their binding mechanisms. We hope that this work will provide significant insights into the design of potent SARS-CoV-2 S protein inhibitors in the future.

2. Materials and methods

2.1 Molecular dynamics (MD) simulation. 2 SARS-CoV-2 S protein complexed with ACE2 was obtained from homology modelling as the initial structures of MD simulations, using the 3D structures of SARS-CoV S protein bound with ACE2 that were downloaded from protein data bank[15] (PDB IDs: 6ACG[6] and 6ACK[6]) as templates[16]. Each simulation system was solvated in a cubic box of TIP3P water extended by 9 Å from the solute, with a rational number of counter ions of Na^+ or Cl^- to neutralize the system. AMBER99SB*-ILDNP[17] force field was used to parameterize the protein. To remove bad contacts formed during the system preparation, 10,000 steps of minimization with constraints (10 kcal/mol/$Å^2$) on heavy atoms, including 5,000 steps of steepest descent minimization and 5,000 steps of conjugate gradient minimization, was performed. Then each system was heated to 300 K within 0.2 ns followed by 0.1 ns equilibration in NPT ensemble. Finally, 5 ns MD simulation on each system at 300 K was performed. The minimization, heating and equilibrium are performed with *sander* program in Amber16. The 5 ns production run was performed with *pmemd.cuda*.

2.2 Binding free energy calculation. To evaluated the binding free energy between the S protein of SARS-CoV and SARS-CoV-2 and ACE2, Molecular Mechanics/Generalized Born Surface Area (MM/GBSA)[18-19] was used to calculated the binding free energy (ΔG) based on 5 ns MD trajectories. In the MM/GBSA, the ΔG was calculated according to equation (1),

$$\Delta G = \Delta H - T\Delta S = \Delta E_{ele} + \Delta E_{VDW} + \Delta G_{gb} + \Delta G_{np} - T\Delta S \quad (1)$$

where ΔE_{ele} and ΔE_{VDW} are the electrostatic and van der Waals energy terms, respectively. ΔG_{gb} and ΔG_{np} are the polar and non-polar solvation free energies, respectively. *Nmode* module in Amber16 was used to calculate the c onformational entropy ($T\Delta S$). In this study, the dielectric constants for solvent and solute were set to 80.0 and 1.0, respectively, and OBC solvation model (igb = 5 and PBradii = 5)[20] was applied. Other parameters are set to default values.

2.3 Conformational change pathway prediction. The up-down conformational change of SARS-CoV-2 S protein was generated by normal model analysis, of which the details have been described in our previous study.[21] Briefly, many iterations of NMA was run to predicted the conformational changes from the initial structures to final target structures gradually. For example, the intermediate structure $R^{(k)}$ in iteration k, is generated by the equation 2 based on the structure $R^{(k-1)}$ in the iteration (k-1):

$$R^{(k)} = R^{(k-1)} + v^{(k)} = R^{(k-1)} + S^{(k)} \sum_{i}^{m^{(k)}} \left(d^{(k-1)} \cdot u_i^{(k)} \right) u_i^{(k)} \quad (2)$$

where $v^{(k)}$ is the displacement combined with $m^{(k)}$ low-frequency eigenmodes that are calculated by NMA. For the i_{th} eigenmode, its displacement is proportional to the projection $d^{(k-1)} u_i^{(k)}$ where $d^{(k-1)}$ is the instantaneous distance vector on eigenvector $u_i^{(k)}$, and scaled by the step size $S^{(k)}$. In this study, the step size is set at 10.0, consisting with our previous study[21]. The starting and final structures are obtained from homology modelling based on the 3D structures in 5X58 and 5X5B corresponding to RBD "down" and "up" state respectively, chosen from Table 1 with the best resolution.

3. Results

3.1 Overview of the SARS-CoV S trimer's structures in the PDB. Amino acid sequence alignment revealed that the S protein of SARS-CoV-2 shares 76% similarity with that of SARS-CoV (Figure 1). The SARS-CoV S protein adopts a homotrimer architecture, of which the RBD undergoes hinge-like conformational switch from prefusion to postfusion. As shown in Table 1, in the PDB, 5 ACE2-free SARS-CoV S trimers are found with three "down" RBDs, which was not observed in any of the ACE2-bound conformations. 4 SARS-CoV S trimers complexed with ACE2 could be found so far (PDB ID: 6ACG, 6ACJ, 6ACK, and 6CS2), of which each a single RBD is in the "up" conformation with different "up" angles ranging from 54.8° to 84.6°, revealing the flexibility of the "up" RBD domain.

```
SARS-CoV-2    MFVFLVLLPLVSSQCVNLTT--RTQLPPAY--TNSFTRGVYYPDKVFRSSVLHSTQDLFL    56
SARS-CoV      MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFL    60
              **:**::* *.*.. ::  *      * *  .*  *******::***..*: *****

SARS-CoV-2    PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQS    116
SARS-CoV      PFYSNVTGFHTIN--------HTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQS    113
              **:**** **:*:       : *.***:**:**:***:*****::***:**:*::.*:**

SARS-CoV-2    LLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFL    176
SARS-CoV      VIIINNSTNVVIRACNFELCDNPFFAVSKPMGT----QTHTMIFDNAFNCTFEYISDAFS    169
              ::*:**:*****:.*:*::*::**:.*    ..    ::.  :::.* ******:*: *

SARS-CoV-2    MDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINIT    236
SARS-CoV      LDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINIT    229
              :*:. *.****:******** **:: :*. : **:****:.**.:*:*:..**:*****

SARS-CoV-2    RFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPL    296
SARS-CoV      NFRAILTAFS------PAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPL    283
              .*:::*: .       :.. * :.****:****:* **:***:***********: :**

SARS-CoV-2    SETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRK    356
SARS-CoV      AELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERK    343
              :* **:*** ::***********: *: .:***************************:* ***** :**

SARS-CoV-2    RISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTG    416
SARS-CoV      KISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTG    403
              :************** ********* ********:********::**:**********

SARS-CoV-2    KIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAG    476
SARS-CoV      VIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPD    463
              ************ ***:***:.*:*:.  ***** ** :*:.:*:*******.  :. .

SARS-CoV-2    STPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKN    536
SARS-CoV      GKPCTP-PALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGPKLSTDLIKN    522
              ..**.    .:***:**:.***  *.*:*****************:********* **:*:**

SARS-CoV-2    KCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS    596
SARS-CoV      QCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCSFGGVS    582
              :*****************  *.*:* ********:: * **:****:* ***** :********

SARS-CoV-2    VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHV    656
SARS-CoV      VITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHV    642
              *******:*. :***********:* .**********.***.****.***** :**********

SARS-CoV-2    NNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPT    716
SARS-CoV      DTSYECDIPIGAGICASYHTVSL----LRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPT    698
              :.***************** * :     **.:.:**:********.*:*****.*****

SARS-CoV-2    NFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK    776
SARS-CoV      NFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDR    758
              **:**:***::****:******.**********:***************** ***:*.***:

SARS-CoV-2    NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ    836
SARS-CoV      NTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQ    818
              **:********:*****:* :**************.**:.****************:**

SARS-CoV-2    YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQI    896
SARS-CoV      YGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQI    878
              **:****** ***********************:*** **:.** :.**.* *:**********

SARS-CoV-2    PFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNA    956
SARS-CoV      PFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNA    938
              ************************* ******.**.:**:.:**:.:.*:**********

SARS-CoV-2    QALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAA    1016
SARS-CoV      QALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAA    998
              ***********************************************************

SARS-CoV-2    EIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFT    1076
SARS-CoV      EIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFT    1058
              **************************************:*************:**:*:**

SARS-CoV-2    TAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNT    1136
SARS-CoV      TAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNT    1118
              *******:***:******** ***.**:*****:.***************:********:***

SARS-CoV-2    VYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNES    1196
SARS-CoV      VYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNES    1178
              ***********************************************************

SARS-CoV-2    LIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKF    1256
SARS-CoV      LIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKF    1238
              ******************:*****************:********.*********

SARS-CoV-2    DEDDSEPVLKGVKLHYT        1273
SARS-CoV      DEDDSEPVLKGVKLHYT        1255
              ****************
```

Figure 1. Sequence alignment of SARS-CoV-1 S protein, and SARS-CoV S protein. Identical residues are denoted by an "*" beneath the consensus position. RBD domain are colored by blue.

Table 1. Summary of SARS-CoV S trimers in the PDB.

PDB ID	Resolution(Å)	Chain	Ligand	RBD states	"up" angle (°)[a]
5WRG[12]	4.3	A	-	down	30.1
		B	-	down	30.1
		C	-	down	30.1
5X58[14]	3.2	A	-	down	31.6
		B	-	down	31.6
		C	-	down	30.7
5X5B[14]	3.7	A	-	up	84.8
		B	-	down	30.9
		C	-	down	30.9
5XLR[12]	3.8	A	-	down	32.1
		B	-	down	32.1
		C	-	down	32.1
6ACC[6]	3.6	A	-	down	33.4
		B	-	down	33.4
		C	-	down	33.4
6ACD[6]	3.9	A	-	down	32.8
		B	-	down	32.8
		C	-	down	32.9
6ACG[6]	5.4	A	-	down	32.6
		B	-	down	32.7
		C	ACE2	up	54.8
6ACJ[6]	4.2	A	-	down	33.0
		B	-	down	33.3
		C	ACE2	up	68.3
6ACK[6]	4.5	A	-	down	33.1
		B	-	down	33.8
		C	ACE2	up	84.6
6CRV[22]	3.2	A	-	-	-
		B	-	-	-
		C	-	-	-
6CRW[22]	3.9	A	-	down	34.3
		B	-	up	68.8
		C	-	down	34.2
6CRX[22]	3.9	A	-	up	71.6
		B	-	up	70.6
		C	-	down	38.1
6CRZ[22]	3.3	A	-	down	34.1
		B	-	up	68.8
		C	-	down	34.1
6CS0[22]	3.8	A	-	down	34.2
		B	-	up	68.8

		C	-	down	34.1
		A	-	up	71.6
6CS1[22]	4.6	B	-	up	70.7
		C	-	down	38.1
		A	-	-	-
6CS2[22]	4.4	B	ACE2	up	74.0
		C	-	-	-
		A	-	down	30.7
6NB6[13]	4.2	B	-	up	77.9
		C	-	up	55.2
		A	-	up	75.3
6NB7[13]	4.5	B	-	up	70.6
		C	-	up	78.5

[a]: The RBD domain "up" angle is determined by residues D392-T608-V972 in SARS-CoV S protein.

3.2 Higher affinity of SARS-CoV-2 S binding to ACE2 than SARS-CoV S. In order to compare the binding affinity of S protein binding to ACE2 between SARS-CoV-2 and SARS-CoV, the MM/GBSA method was used to predict the binding free energy, which has been recommended with more accurate prediction than many empirical scoring functions applied in protein-protein docking.[23] The starting structures of SARS-CoV-2 S protein complexed with ACE2 were obtained from homology modelling using 6ACG and 6ACK as templates, chosen from ACE2 bound SARS-CoV S protein in Table 1 with the largest and smallest "up" angles, respectively.

As shown in Table 2, in the results of simulations started from conformation of 6ACG, the calculated binding free energies of SARS-CoV-2 S binding to ACE2 is -21.74±0.65 kcal/mol, which is obviously stronger than that of SARS-CoV S protein complexed with ACE2 (-10.17±0.63 kcal/mol). It provides an evidence that the SARS-CoV-2 S binds ACE2 with higher affinity than SARS-CoV S, which is one of the reasons of the fact that SARS-CoV-2 is more readily transmitted from human-to-human than SARS-CoV, being in good agreement with the experimental results[24]. In addition, in the results of simulations started from conformation of 6ACK, the calculated binding free energies of SARS-CoV-2 S binding to ACE2 (--29.90±0.80 kcal/mol) is also stronger than that of SARS-CoV S binding to ACE2 (-15.46±0.68 kcal/mol), revealing that the SARS-CoV-2 S protein could maintain higher affinity binds to ACE2 even though the flexible "up" RBD. One can also conclude that the SARS-CoV-2 S protein has higher affinity with more "up" RBD domain, according to calculated binding free energies based on the simulations started by structures modelled from 6ACG and 6ACK, with RBD "up" angles of 54.8° to 84.6°, respectively.

Table 2. Components of the binding free energy (kcal/mol) calculated by MM/GBSA

method*

Energy term	6ACG ("up" angle = 54.8)		6ACK ("up" angle = 84.6)	
	SARS-CoV	SARS-CoV-2	SARS-CoV	2019-nCoV
E_{vdw}	-80.57±0.46	-87.07±0.49	-96.89±0.59	-105.05±0.36
E_{ele}	65.07±0.58	-673.99±3.96	-7.57±0.32	-641.25±4.07
E_{gb}	0.90±0.02	737.98±3.86	83.60±0.26	714.56±3.65
E_{np}	-10.31±0.06	-12.21±0.06	-12.93±0.08	-15.03±0.07
ΔH	-24.91±0.50	-35.30±0.60	-33.80±0.74	-46.77±0.61
$-T\Delta S$	-14.74±0.76	-13.56±0.70	-18.34±0.62	-16.87±0.98
ΔG	-10.17±0.63	-21.74±0.65	-15.46±0.68	-29.90±0.80

*: The statistical error was estimated based on 0.5-5 ns MD simulation trajectory. 500 snapshots evenly extracted from the 0.5-5 ns MD trajectory of complex were used for MM/GBSA calculations and 10 snapshots for the entropy term calculations.

3.3 Comparison of the structure-affinity relationships between SARS-CoV-2 S and SARS-CoV S. To identify key residues in the S-ACE2 interactions, the binding free energies were decomposed into residues by the *MMPBSA.py* module in Amber 16. As shown in Figure 2, the interaction profiles are somewhat similar between SARS-CoV and SARS-CoV-2. For example, in the results of simulations started by 6ACG, residues Y442, L443, P462, L472, N473, Y475, Y484, T487, and Y491 are favorable energy contributors in SARS-CoV S protein bound with ACE2, which are corresponding to residues L455, F456, A475, F486, N487, Y489, Q498, N501, and Y505 in SARS-CoV-2 S protein by sequence alignment, respectively (Figure 2A). In particular, the residue Y491 contributes -4.03±0.60 kcal/mol in SARS-CoV S protein, and the corresponding residue Y505 contributes -4.23±0.56 kcal/mol in SARS-CoV-2 S protein. The difference is extra favorable energy contributors Y449, Q493, G496, T500, and G502 in SARS-CoV-2 S protein, especially for residue Q493 (-3.49±0.48 kcal/mol), suggesting more interactions formed in SARS-CoV-2 S protein binding to ACE2, which accounts for the higher binding affinity of SARS-CoV-2 S protein than that of SARS-CoV S protein. Similarly, in the results of simulations started by 6ACK, there are more residues whose energy contribution more than 1.0 kcal/mol in the S-ACE2 interface, even though the binding affinity of SARS-CoV S protein is higher that of simulation started by 6ACG (Figure 2B).
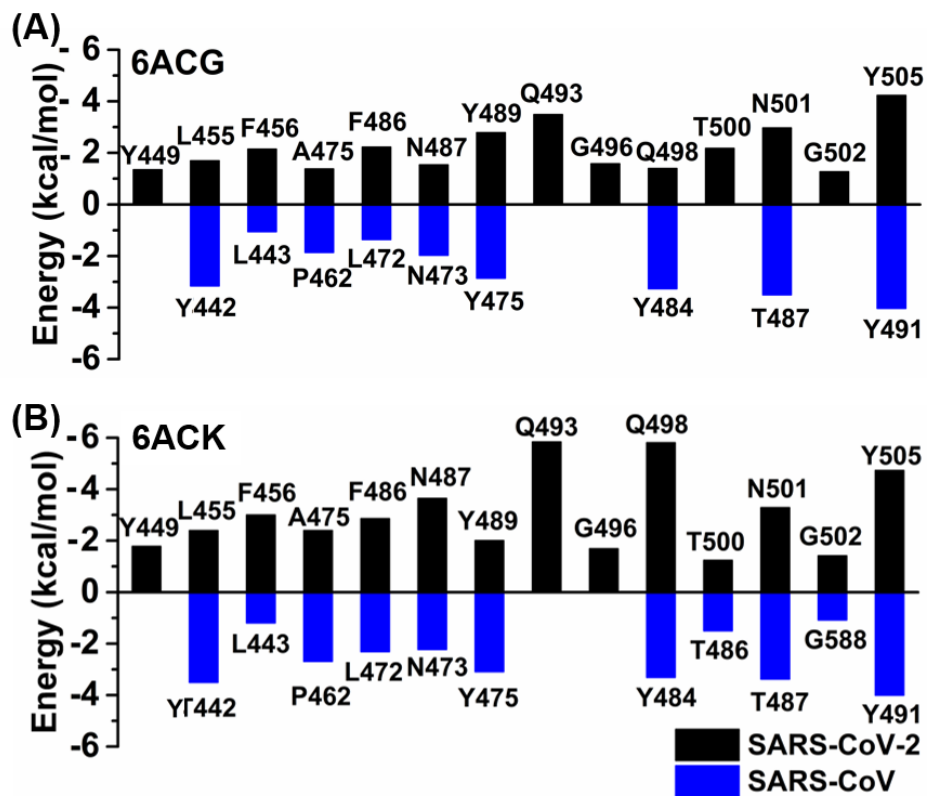
Figure 2. S protein Residue-ACE2 interaction spectrum of SARS-CoV-2 (colored black) and SARS-CoV (colored blue). The initial structures of MD simulation were based on the 3D structures of 6ACG (A) and 6ACK (B). The residues that contribute less than -1.00 kcal/mol to binding energy were labeled in the black fonts.

3.4 Identification of the ACE2-accessible RBD "up" angle of SARS-CoV-2 S. The two states, "down" and "up" conformations, correspond to the receptor-inaccessible and receptor-accessible states, respectively. However, as shown in Table 1, ACE2-bound SARS-CoV S protein still have different RBD "up" angles, suggesting that the RBD should "up" to a receptor-accessible angles before binding to ACE2. To identify the ACE2-accessible RBD "up" angle, we calculated atomic-level "down" to "up" conformational change of SARS-CoV-2 S protein by normal modes analysis (Figure 3A), starting by "down" conformation modelled by 5X58 chosen from Table 1 with the best resolution. By aligning the RBD-ACE2 complexes of 6ACG with conformations along the conformational change pathway, we found that only the RBD "up" to 52.2°, there is no atomic collision between ACE2 and S protein, being in well agreement with experimental results (Figure 3B). For examples, as shown in Table 1, all the "down" conformations of SARS-CoV S protein have RBD "up" angle less than 52.2°. In addition, the smallest RBD "up" angle 54.8° in 6ACG is still larger than 52.2°.
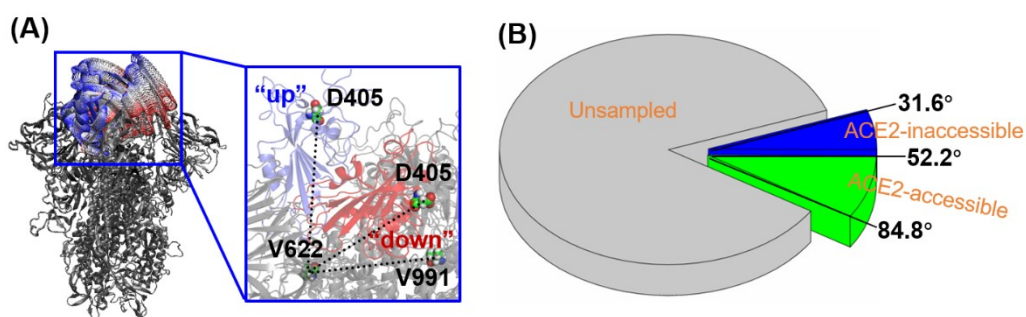
Figure 3. (A), Conformational change pathway of SARS-CoV-2 S protein generated by NMA. The "up" angle is determined by residues D405-V622-V991, corresponding to residues D392-T608-V972 in SARS-CoV S protein. (B), The RBD "up" angle of the ACE2-inaccessible (blue), ACE2-accessible (green), and unsampled (gray) conformations for the SARS-CoV-2 S.

4. Conclusions

The outbreak of the SARS-CoV-2 has seriously threatened the global health, which caused more than 2,000 deaths in China as of 27 January 2020. However, there is no approved effective drug. The SARS-CoV-2 spike (S) glycoprotein is a key target for drug discovery and design, due to its indispensable function for viral infection and fusion by using human angiotensin-converting enzyme 2 (ACE2) as a receptor. To facilitate the development of inhibitor to S-ACE2 interactions, we used various computational techniques to study the binding mechanisms of S-ACE2. Compared with SARS-CoV, SARS-CoV-2 S protein has obvious higher affinity binds to ACE2 predicted by MM/GBSA, which might account for the ease of transmission from human-to-human of SARS-CoV-2. The binding free energy decomposition analysis further showed that more interactions formed in SARS-CoV-2 S protein binding to ACE2 accounts for the higher binding affinity. In addition, from the binding free energies of SARS-CoV-2 S proteins with different RBD "up" angle, it could be found that SARS-CoV-2 S protein has higher affinity binds to ACE2 with more "up" RBD. Therefore, to identify an ACE2-accessible RBD "up" angle, the "down" to "up" conformational change of SARS-CoV-2 S protein was generated by NMA. The results suggested that 52.2° is an ACE2-accessible RBD "up" angle, being consistent with experimental results, which also suggested that conformations between RBD "down" and up to 52.2° is ideal target structures for SARS-CoV-2 S inhibitor to its conformational change. We hope that this work will provide significant insights into the design of potent SARS-CoV-2 S protein inhibitors to address the ongoing public health crisis.

References

1.   Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.;

Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W., A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020**.

2. Zhou, P.; Yang, XL.; Wang, XG.; Hu, B.; Zhang, L.; Zhang, W.; Si, HR.; Zhu, Y.; Li, B.; Huang, CL.; Chen, HD.; Chen, J.; Luo, Y.; Guo, H.; Jiang, RD.; Liu, MQ.; Chen, Y.; Shen, XR.; Wang, X.; Zheng, XS.; Zhao, K.; Chen, QJ.; Deng, F.; Liu, LL.; Yan, B.; Zhan, FX.; Wang, YY.; Xiao, GF.; Shi, ZL., A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**.

3. Wu, F.; Zhao, S.; Yu, B.; Chen, YM.; Wang, W.; Song, ZG.; Hu, Y.; Tao, ZW.; Tian, JH.; Pei, YY.; Yuan, ML.; Zhang, YL.; Dai, FH.; Liu, Y.; Wang, QM.; Zheng, JJ.; Xu, L.; Holmes, E. C.; Zhang, YZ., A new coronavirus associated with human respiratory disease in China. *Nature* **2020**.

4. Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; Qiu, Y.; Wang, J.; Liu, Y.; Wei, Y.; Xia, J.; Yu, T.; Zhang, X.; Zhang, L., Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **2020,** *395*, 507-513.

5. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K. S. M.; Lau, E. H. Y.; Wong, J. Y.; Xing, X.; Xiang, N.; Wu, Y.; Li, C.; Chen, Q.; Li, D.; Liu, T.; Zhao, J.; Li, M.; Tu, W.; Chen, C.; Jin, L.; Yang, R.; Wang, Q.; Zhou, S.; Wang, R.; Liu, H.; Luo, Y.; Liu, Y.; Shao, G.; Li, H.; Tao, Z.; Yang, Y.; Deng, Z.; Liu, B.; Ma, Z.; Zhang, Y.; Shi, G.; Lam, T. T. Y.; Wu, J. T. K.; Gao, G. F.; Cowling, B. J.; Yang, B.; Leung, G. M.; Feng, Z., Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **2020**.

6. Song, W.; Gui, M.; Wang, X.; Xiang, Y., Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog* **2018,** *14* (8), e1007236.

7. Li, F., Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* **2016,** *3* (1), 237-261.

8. Bosch, B. J.; van der Zee, R.; de Haan, C. A.; Rottier, P. J., The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J. Virol.* **2003,** *77* (16), 8801-11.

9. Walls, A. C.; Tortorici, M. A.; Snijder, J.; Xiong, X.; Bosch, B. J.; Rey, F. A.; Veesler, D., Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** *114* (42), 11157-11162.

10. Belouzard, S.; Millet, J. K.; Licitra, B. N.; Whittaker, G. R., Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. *Viruses* **2012,** *4* (6), 1011-1033.

11. Jesper, P.; Nianshuang W.; Kizzmekia S. C.; Daniel W.; Robert N. K.; Hannah L. T.; Christopher A. C.; Michelle M. B.; L. W.; Wei S.; Konge W.; L. Erica; A. N. Kettenbachb; Mark R. Denisond; James D. Chappelld; Barney S. Grahamc; Andrew B. Warda; Jason S. McLellanb, Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci. U S A* **2017,** *1707304114*, 7348-7357.

12. Gui, M.; Song, W.; Zhou, H.; Xu, J.; Chen, S.; Xiang, Y.; Wang, X., Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res.* **2017,** *27* (1), 119-129.

13. Walls, A. C.; Xiong, X.; Park, Y. J.; Tortorici, M. A.; Snijder, J.; Quispe, J.; Cameroni, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.; Zambon, M.; Rey, F. A.; Corti, D.; Veesler, D., Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **2019,** *176* (5), 1026-1039 e15.

14. Yuan, Y.; Cao, D.; Zhang, Y.; Ma, J.; Qi, J.; Wang, Q.; Lu, G.; Wu, Y.; Yan, J.; Shi, Y.; Zhang, X.; Gao, G. F., Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat. Communications* **2017,** *8* (1).

15.   Burley, S. K.; Berman, H. M.; Bhikadiya, C.; Bi, C.; Chen, L.; Di Costanzo, L.; Christie, C.; Dalenberg, K.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ghosh, S.; Goodsell, D. S.; Green, R. K.; Guranović, V.; Guzenko, D.; Hudson, B. P.; Kalro, T.; Liang, Y.; Lowe, R.; Namkoong, H.; Peisach, E.; Periskova, I.; Prlić, A.; Randle, C.; Rose, A.; Rose, P.; Sala, R.; Sekharan, M.; Shao, C.; Tan, L.; Tao, Y.-P.; Valasatava, Y.; Voigt, M.; Westbrook, J.; Woo, J.; Yang, H.; Young, J.; Zhuravleva, M.; Zardecki, C., RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic. Acids Res.* **2018,** *47* (D1), D464-D474.

16.   Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic. Acids Res.* **2018,** *46* (W1), W296-W303.

17.   Aliev, A. E.; Kulke, M.; Khaneja, H. S.; Chudasama, V.; Sheppard, T. D.; Lanigan, R. M., Motional timescale predictions by molecular dynamics simulations: case study using proline and hydroxyproline sidechain dynamics. *Proteins* **2014,** *82* (2), 195-215.

18.   Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., 3rd, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000,** *33* (12), 889-97.

19.   Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A., Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate−DNA Helices. *J. Am. Chem. Soc.* **1998,** *120* (37), 9401-9409.

20.   Onufriev, A.; Bashford, D.; Case, D. A., Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004,** *55* (2), 383-94.

21.   Wang, J.; Shao, Q.; Xu, Z.; Liu, Y.; Yang, Z.; Cossins, B. P.; Jiang, H.; Chen, K.; Shi, J.; Zhu, W., Exploring Transition Pathway and Free-Energy Profile of Large-Scale Protein Conformational Change by Combining Normal Mode Analysis and Umbrella Sampling Molecular Dynamics. *J. Phys. Chem. B* **2013,** *118* (1), 134-143.

22.   Kirchdoerfer, R. N.; Wang, N.; Pallesen, J.; Wrapp, D.; Turner, H. L.; Cottrell, C. A.; Corbett, K. S.; Graham, B. S.; McLellan, J. S.; Ward, A. B., Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci. Rep.* **2018,** *8* (1), 15701.

23.   Chen, F.; Liu, H.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T., Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking. *Phys. Chem. Chem. Phys.* **2016,** *18* (32), 22129-22139.

24.   Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.; Graham, B. S.; McLellan, J. S., Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation. *Sciences* **2020,** *eabb2507*.